Bargaining in Dynamic Games

Leon A. Petrosyan

St. Petersburg University

26^{th} European Conference on Operational Research

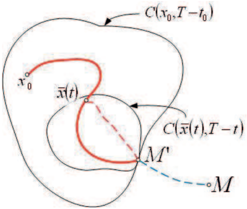
Rome 1-4 July, 2013

1. Classical control problem.

$$\dot{x} = f(x, u), \quad x \in R^n, u \in U \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$
$$H(x(T)) = -\rho(x(T), M).$$

$C(x_0, T - t_0)$ – reachability set.
$\bar{x}(t)$ – optimal trajectory.

$\Gamma(x_0, T - t_0), \quad \Gamma(\bar{x}(t), T - t), \quad C(\bar{x}(t), T - t)$



R. Bellmann
Time-consistency, Strong Time-consistency.

2. Multycriterial control.

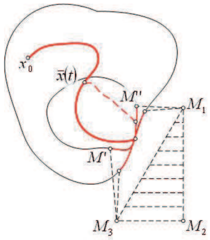
$$\dot{x} = f(x, u), \quad x \in R^n, u \in U \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$
$$H(x(T)) = \{H_1(x(T)), \ldots, H_k(x(T))\}.$$

Let $k = 3$, $H_i(x(T)) = -\rho(x(T), M_i)$.

Pareto-optimal solution.

$\bar{x}(t)$ – Pareto-optimal trajectory.

$\Gamma(x_0, T - t_0)$, $\Gamma(\bar{x}(t), T - t)$, $C(x_0, T - t_0)$, $C(\bar{x}(t), T - t)$, $P(x_0, T - t_0)$, $P(\bar{x}(t), T - t)$



TC but not STC

3. Nash bargaining solution in Differential Games.

$$\dot{x} = f(x, u_1, \ldots, u_n), \quad x \in R^m, u_i \in U_i \subset CompR^l,$$
$$x(t_0) = x_0, \quad t \in [t_0, T],$$

The payoff of player $i \to H_i(x(T))$,
$\Gamma(x_0, T - t_0)$
$W(x_0, T - t_0; \{i\})$ – the garanteed payoff of player i

NB.

$$\max_{x' \in C(x_0, T - t_0)} \prod_{i=1}^{n} (H_i(x') - W(x_0, T - t_0; \{i\})) = \prod_{i=1}^{n} (H_i(\bar{\bar{x}}) - W(x_0, T - t_0; \{i\}))$$

$$\bar{x}(t), \quad x_0 \to \bar{\bar{x}}, \quad \Gamma(\bar{x}(t), T - t), t \in [t_0, T - t_0], \quad W(\bar{x}(t), T - t; \{i\})$$

$$\max_{x' \in C(\bar{x}(t), T - t)} \prod_{i=1}^{n} (H_i(x') - W(\bar{x}(t), T - t; \{i\})) = \prod_{i=1}^{n} (H_i(\bar{\bar{x}}(\bar{x}(t))) - W(\bar{x}(t), T - t; \{i\}))$$

$\bar{\bar{x}}(\bar{x}(t)) \neq const \neq \bar{x}$
NB, not TC, not STC

Regularization. Here we consider simple case of the game with terminal payoff and prescribed duration.

Denote the payoff function as

$$K_i(x_0, T - t_0; u_1(\cdot), \dots, u_n(\cdot)) = H_i(x(T)),$$

where $x(t)$ is the solution of problem

$$\dot{x} = f(x, u_1, \dots, u_n) \tag{1}$$

$$x(t_0) = x_0$$

Denote the corresponding differential game by $\Gamma(x_0, T - t_0)$. Consider a Bargaining Problem for each subgame $\Gamma(y, T - t)$, where $y \in C(x_0, t)$, and $C(x_0, t)$ is the reachability set for system (1). We can have in mind as bargaining solution Nash Bargaining (NB), (KS). Suppose we define for each $y \in C(x_0, t)$, $t \in [t_0, T]$ the NB solution and denote it by $D(y, T - t)$.

$$D(y, T - t) = \{H_i(\widetilde{x}(T; y, t))\}$$

Suppose $\widetilde{x}(t)$, $t \in [t_0, T]$ is the bargain trajectory in the game $\Gamma(x_0, T - t_0)$ leading from initial point x_0 to bargaining point $\widetilde{x}(T; x_0, t_0)$. As we mention the NB solution as many other bargaing solutions (including Kalai–Smorodinsky) are time inconsistent, which means that

$$D(\widetilde{x}(t), T - t) \neq const \qquad t \in [t_0, T].$$

This makes difficult the implementation of NB in practice.

Here we purpose the regularization mechanism similar to IDP we used considering the similar time-inconsistency problem in differential cooperative games. To look different from IDP (imputation distribution procedure) we shall call in PDP (payoff distribution procedure).

Introduce the function $g_i(\tau)$, $\tau \in [t_0, T]$ such that

$$D_i(x_0, T - t_0) = H(\widetilde{x}(T; x_0, T - t_0)) = \int_{t_0}^{T} g_i(\tau) d\tau.$$

$g_i(\tau)$ is a new control variable ($\tau \in [t_0, T]$) used to distribute the payoff over the time interval $[t_0, T]$.
Time-consistency condition

$$D_i(x_0, T - t_0) = \int_{t_0}^{t} g_i(\tau) d\tau + D_i(\widetilde{x}(\tau), T - \tau).$$

If function D_i is differentiable we get

$$g_i(t) = -\frac{d}{dt} D_i(\widetilde{x}(\tau), T - t),$$

or

$$g_i(t) = -\sum_{i=1}^{n} \frac{\partial D_i}{\partial x_i} f_i(\widetilde{x}(t), \widetilde{u}_1(t), \dots, \widetilde{u}_n(t)), \qquad i = 1, \dots, n.$$

Kalai–Smorodinsky solution

$$\dot{x} = f(x, u_1, \dots, u_n), \qquad x \in R^m, u_i \in U \subset CompR^l$$

$$x(t_0) = x_0, \qquad t \in [t_0, T].$$

The payoff of player i – $H_i(x(T))$
$\Gamma(x_0, T - t_0)$
$W(x_0, T - t_0; \{i\})$ – the max guaranteed payoff of player i
KS–solution
Consider the set $\aleph(x_0, T - t_0)$ of all possible payoffs of players.

$$\aleph(x_0, T - t_0) = \{H_1(x(T)), \dots, H_n(x(T))\}$$

for all possible $x(T) \in C(x_0, T - t_0)$ where $C(x_0, T - t_0)$ is the reachability set of (1).
Correspondingly denote by $\aleph(y, T - t)$, $y \in C(x_0, T - t_0)$ the set of all possible payoffs of players in subgames starting from y with duration $T - t$ ($\Gamma(y, T - t)$).
Denote by

$$\overline{H}_i = \max_{x(T) \in C(x_0, T - t_0)} H(x(T)).$$

Consider a line L connecting the "status point" W with $\overline{H}$.

The intersection of L and $\aleph(x_0, T - t_0)$ is called Kalai–Smorodinsky solution (KS–solution).

Example. KS–solution is time–inconsistent

$$\dot{x} = u_1 + u_2, \qquad |u_1| \leq 1, |u_2| \leq 1, x \in R^2, u_1, u_2 \in R^2$$
$$x_0 = (6, 3), t \in [0, 2], x = (x_1, x_2)$$
$$H_i(x_0, 2; u_1, u_2) = -x(2), K_2(z_0, 2; u_1, u_2) = -|y(2)|$$

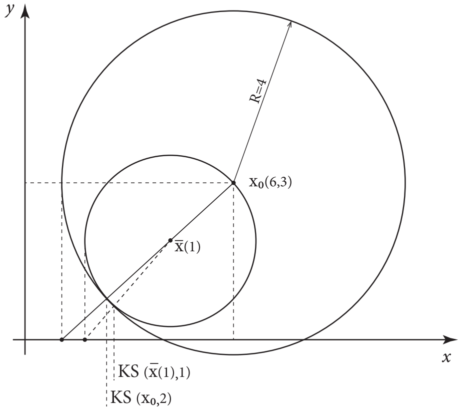
Problem $\Gamma(x_0, 2)$
$\underline{W}(x_0, 2) = (-6, 3)$
$\overline{H}(x_0, 2) = (-2, 0)$
and corresponds to the point $\widehat{x}(x_0, 2) = (2, 0)$, since $\max H_1 = -2$, $\max H_2 = 0$.
$C^2(6, 3; 2)$ is circle with center $(6, 3)$ and radius 4.

Optimal trajectory is to move from $x_0 = (6, 3)$ to $\widehat{x}(x_0, 2) = (2, 0)$ with maximal velocity till the intersection with $C(6, 3; 2)$. The intersection point is KS–solution of $\Gamma(x_0, 2)$.

We see that when moving along KS–solution trajectory $\overline{x}(t)$ connecting x_0 with KS–solution of the problem $\Gamma(x_0, 2)$ in subgame $\Gamma(\overline{x}(1), 1)$ we get a new KS–solution which do not coincide with previous KS–solution of the game $\Gamma(x_0, 2)$.

$$KS(x_0, 2) \neq KS(\overline{x}(1), 1).$$

(time–inconsistency)



4. Differential Cooperative Game.

Differential Cooperative Game $\Gamma(x_0, T - t_0)$ with prescribed duration $T - t_0$ from the initial position x_0.

$$\dot{x} = f(x, u_1, \ldots, u_n), \quad x \in R^n, u_i \in U_i \tag{1}$$

integral payoff

$$K_i(x_0, T - t_0; u_1, \ldots, u_n) = \int_{t_0}^{T} h_i(x(t))dt, \quad h_i > 0, i = 1, \ldots, n.$$

Cooperative form of $\Gamma(x_0, T - t_0)$.

Cooperative behavior $u^*(t) = \{u_1^*(t), \ldots, u_n^*(t)\}$

$$\begin{aligned} \sum_{i=1}^{n} K_i(x_0, T - t_0; u_1^*, \ldots, u_n^*) &= \\ &= \max_{u_1, \ldots, u_n} \sum_{i=1}^{n} K_i(x_0, T - t_0; u_1, \ldots, u_n) = \\ &= \sum_{i=1}^{n} \int_{t_0}^{T} h_i(x^*(t))dt = v(N; x_0, T - t_0), \end{aligned}$$

$x^*(t)$ – cooperative trajectory.

Characteristic Function in $\Gamma(x_0, T - t_0)$.

$v(S; x_0, T - t_0), \quad S \subset N,$

superadditivity: $v(S_1 \cup S_2; x_0, T - t_0) \geq v(S_1; x_0, T - t_0) + v(S_2; x_0, T - t_0)$, $S_1 \cap S_2 = \emptyset$.

There are different ways on how to define c. f.

a. Classical: $v(S; x_0, T - t_0) = Val\Gamma_{S,N\setminus S}(x_0, T - t_0)$, where $\Gamma_{S,N\setminus S}(x_0, T - t_0)$ is a zero-sum game played upon the structure of game $\Gamma(x_0, T - t_0)$ between S as player 1 and $N \setminus S$ as player 2.

b. $v(S; x_0, T - t_0) = \sum_{i \in S} K_i(x_0, T - t_0; \bar{u}_S, \bar{u}_{N\setminus S})$, where $(\bar{u}_S, \bar{u}_{N\setminus S})$ is some giver NE in $\Gamma'_{S,N\setminus S}$ played as non zero-sum game over the structure of $\Gamma(x_0, T - t_0)$ between two players: coalition S as player 1 and $N \setminus S$ as player 2

c. $v(S; x_0, T - t_0) = \max_{u_S = \{u_i, i \in S\}} \sum_{i \in S} K_i(x_0, T - t_0; \bar{u}||u_S)$, where $\bar{u} = (\bar{u}_1, \dots, \bar{u}_n)$ is some fixed NE in $\Gamma(x_0, T - t_0)$.

L. Petrosjan, G. Zaccour Time-consistent Shapley value allocation of pollution cost reduction // Journal of Economics Dunamics & Control, 27 (2003), pp. 381-398.

Let $E(x_0, T - t_0)$ be the imputation set in $\Gamma(x_0, T - t_0)$:

$$E(x_0, T - t_0) = \{\xi = (\xi_i) : \sum_{i=1}^{n} \xi_i = v(N; x_0, T - t_0), \xi_i \geq v(\{i\}; x_0, T - t_0), i \in N\}.$$

Denote by $C^{t-t_0}(x_0), t \in [t_0, T]$ reachable set of the (1).
For each $y \in C^{t-t_0}(x_0)$ consider a subgame $\Gamma(y, T - t)$ of the game $\Gamma(x_0, T - t_0)$, with corresponding c. f. $v(S; y, T - t)$ and set of imputations $E(y, T - t)$.

Definition. A point-to-set mapping $C(y, T - t) \subset E(y, T - t)$ defined for all $y \in C^{t-t_0}(x_0), t \in [t_0, T]$ is call *solution concept* (SC) in the family of subgames $\Gamma(y, T - t)$.

In special cases $C(y, T - t)$ may be a core, NM-solution, Shapley value, nucleous etc.

What happens when the game develops along the cooperative trajectory $x^*(t)$?
We pass through current subgame $\Gamma(x^*(t), T - t)$, wilingly or not updating the current SC $\leftrightarrow C(x^*(t), T - t)$.

Imputation Distribution Procedure (IDP).

Let $\bar{\xi} \in C(x_0, T - t_0)$ and $\beta_i(t), i \in N, t \in [t_0, T]$ satisfies the condition

$$\bar{\xi} = \int_{t_0}^{T} \beta_i(t)dt, \quad i \in N, \quad \beta_i \geq 0.$$

$\beta_i(t)$ is called IDP.

Define

$$\bar{\xi}(\theta) = \int_{t_0}^{\theta} \beta_i(t)dt, \quad i \in N, \quad \beta_i \geq 0.$$

Definition. The SC $C(x^*(t), T - t), t \in [t_0, T]$ is called *time-consistent* (TC) if there exist such IDP $\beta(t) = \{\beta_i(t)\}$ that

$$\bar{\xi} - \bar{\xi}(\theta) \in C(x^*(\theta), T - \theta)$$

for all $\theta \in [t_0, T]$.

Definition. The SC $C(x^*(t), T - t), t \in [t_0, T]$ is called *strongly time-consistent* (STC) if there exist such IDP $\beta(t) = \{\beta_i(t)\}$ that

$$\bar{\xi}(\theta) \oplus \bar{C}(x^*(\theta), T - \theta) \subset C(x_0, T - t_0)$$

for all $\theta \in [t_0, T]$. Here $\bar{\xi} \oplus A$ means the set of all possible vectors $\bar{\xi} + \eta$ for all $\eta \in A$.

Consider $C(x^*(t), T-t)$ along $x^*(t), t \in [t_0, T]$. Suppose we can construct a differentiable selector $\xi^t \in C(x^*(t), T-t)$, then we can easily get for $\beta(t)$ the following formula

$$\bar{\xi} = \bar{\xi}(\theta) + \xi^t \quad \rightarrow \quad \bar{\xi} = \int_{t_0}^{\theta} \beta_i(t) dt + \xi^t$$

$$\beta_i(t) = -\frac{d}{dt} \xi^t$$

If ξ^t can be chosen as monotonic nonincreasing (which is very possible since $h_i > 0$, then $\beta_i \geq 0$, and SC is TC.
If the case (for instance) $C(y, T-t)$ is a Shapley value, we get

$$\beta_i(t) = -\sum_{S \subset N, S \ni i} \frac{(n-s)!(s-1)!}{n!} \left[\frac{d}{dt} v(x^*(t), T-t; S) - \right.$$

$$\left. - \frac{d}{dt} v(x^*(t), T-t; S \setminus \{i\}) \right]$$

and we need only differentiability of the value function (c. f.) $v(x, T-t; S)$.

References

Neumann J., Morgenstern O. (1947). "Theory of Games and Economic Behavior Princeton.

Nash J. (1951). "Non-cooperative games Ann. Mathematics, 54, pp. 286–295.

Petrosjan L. A. (1993). "Differential Games of Pursuit World Scientific, Singapore.

Petrosyan L. A., Yeung D. W. K. (2006) Cooperative Stochastic Differential Games. Springer, 242 p.

Shapley L.S. (1953). "A Value for *n*-Person Games Contributions to the Theory of Games (in H.W. Kuhn and A.W. Tucker (eds.), Princeton, Princeton University Press, pp. 307–315.